Automatica, Vol. 11, pp. 451–460. Pergamon Press, 1975. Printed in Great Britain

Tracking in a Cluttered Environment With Probabilistic Data Association*

Dépistage dans une Ambiance Encombrée, avec Association Probabilistique des Données

Verfolgung in einer örtlich gestörten Umgebung unter Verwendung von Wahrscheinlichkeitsdaten

YAAKOV BAR-SHALOM† and EDISON TSE†

Tracking a target with uncertainty in the origin of the measurements is accomplished with an algorithm, suitable for real-time implementation, which utilizes the a posteriori *probabilities of the measurements having originated from the target.*

Summary—This paper presents a new approach to the problem of tracking when the source of the measurement data is uncertain. It is assumed that one object of interest ('target') is in track and a number of undesired returns are detected and resolved at a certain time in the neighbourhood of the predicted location of the target's return. A suboptimal estimation procedure that takes into account all the measurements that might have originated from the object in track but does not have growing memory and computational requirements is presented. The probability of each return (lying in a certain neighborhood of the predicted return, called 'validation region') being correct is obtained—this is called 'probabilistic data association' (PDA). The undesired returns are assumed uniformly and independently distributed. The estimation is done by using the PDA method with an appropriately modified tracking filter, called PDAF. Since the computational requirements of the PDAF are only slightly higher than those of the standard filter, the method can be useful for real-time systems. Simulation results obtained for tracking an object in a cluttered environment show the PDAF to give significantly better results than the standard filter currently in use for this type of problem.

1. INTRODUCTION

THIS paper presents a new, suboptimal approach to the problem of tracking when the source of the measurement data is uncertain. This can occur, when a sensor, e.g. a radar, is operating in an environment in which there is clutter or the false-alarm rate is high. It is assumed that undesirable returns occur independently in time and space and that no inference from past data can be made on the location, nature or number of these returns. The reason for this assumption comes from the fact that in many problems of practical interest there is little knowledge about the environment. The dynamics of the target of interest are assumed to be known as well as the covariance of the driving noise; the method presented here does not deal with targets that might maneuver. The motivation of this study came from a realistic problem in which white noise is a good model as input as described, for example, in [1], and where undesirable returns can seriously affect the estimator's performance.

When tracking a target in such an environment there might be several 'candidate returns'. It is assumed that the returns known for sure not to have originated from the target of interest, using, e.g. their signature, were already discarded. However, there are still several returns resolved by the receiver that, based upon the available information, predicted location of the correct return and its statistics and their measured location, might have originated from the object in track. The usual procedure, using a 'standard filter', is to select one of them according to a certain rule, e.g. 'nearest neighbor', and then use it for updating the estimate of the state of the object in track. In this case, even if the dynamics and measurement equations are linear and the noises Gaussian, a linear (Kalman) filter will not be optimal. This is due to the fact that with some non-zero probability, the measurement which was selected for the update did not originate from the object in track. Therefore, the performance of this filter would be degraded because it is too 'optimistic'—it would attach too

* Received 9 July 1973; revised 8 February 1974; revised 23 July 1974; revised 23 December 1974. The original version of this paper was not presented at any IFAC meeting. It was recommended for publication in revised form by associate editor I. Rhodes.

† Systems Control, Inc., Palo Alto, California 94304, U.S.A.

high a confidence in the estimates—and the result might be loss of track. Another method commonly in use in this case is 'track split', i.e. when more than one return is considered to be valid, a new track is formed for every possible path. This method is very conservative and might be infeasible in a real-time system because of being computationally costly.

Sea [2] and Singer *et al.* [3, 4] treated the problem considered here and developed a modified tracking filter. This tracking filter took into account the possibility that the return chosen for the update, the nearest to the predicted measurement, was an incorrect one. However, this was done using only the *a priori* probability that the selected return originated from the object in track. Jaffer and Bar-Shalom [6] derived a modified filter that utilized the *a posteriori* probability that the selected return originated from the object in track. This algorithm incorporated real-time information to modify the Kalman filter, as opposed to [2–4] which used only *a priori* information to this purpose. The above procedures [2–4, 6] are of the nearest-neighbor type with modified filter.

The approach suggested here is to obtain an estimator which incorporates *all* the measurements that might have originated from the object in track rather than selecting at each time one of them as in the previous procedures. This estimator makes use of the *a posteriori* probability that each of the validated measurements originated from the object in track—therefore, it incorporates all the available information. These probabilities are obtained using the locations of the measurements.* Additional information, e.g. target signature, might also be incorporated in these probabilities. In the computation of these probabilities it is assumed that the innovation corresponding to the correct return is normally distributed. While this is only an approximation, it allows us to obtain a fixed memory algorithm. The simulation results indicate that the performance of the PDAF is significantly superior to that of a standard filter; this is believed to justify the practicality of the above approximation. The algorithm is adaptive to the environment in the following sense: the actual number of returns observed and their location will affect the confidence, in terms of the covariance, of the estimate. While it is obvious that no filter is capable of tracking in an arbitrarily dense cluttered environment, the importance of the method presented is that it extends the range, in terms of clutter density, in which one can reliably track targets with reasonable computational requirements.

* A preliminary version of this method appears in [5]. After [5] appeared, Singer *et al.* [7] obtained a filter that was combining track splitting with data correlation using *a posteriori* statistics.

The problem is formulated in Section 2 and the probabilistic data association procedure is described in Section 3. The new filter, called probabilistic-data-association filter (PDAF), is presented in Section 4. Also the computational requirements of the PDAF are shown and compared with those of the track-split filter (TSF). The simulation results presented in Section 5 show a notable improvement in the performance of the PDAF over the nearest-neighbor standard filter for a tracking problem where the use of the TSF would be prohibitive.

2. FORMULATION OF THE PROBLEM

The dynamics of the object in track are modelled by the equation

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{w}_k, \quad k = 0, 1, \ldots, \tag{2.1}$$

where $\mathbf{x}_k$ is the n-dimensional state vector, F_k is the (known) transition matrix and $\mathbf{w}_k$ is the process noise, assumed to be normally distributed with mean zero and known variance

$$E(\mathbf{w}_k \mathbf{w}_j') = \mathbf{Q}_k \delta_{kj} \tag{2.2}$$

The initial state is also normally distributed with mean $\hat{\mathbf{x}}_{0|0}$ and covariance $\mathbf{P}_{0|0}$, independent of $\mathbf{w}_k$.

The measurement system is modelled as follows. If the measurement originates from the object in track, then

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{y}_k, \quad k = 1, \ldots, \tag{2.3}$$

where $\mathbf{H}_k$ is a known $(r \times n)$ matrix and the measurement noise $\mathbf{y}_k$, independent of $\mathbf{w}_j$ and $\mathbf{x}_0$, is normally distributed with mean zero and known variance

$$E\mathbf{v}_k \mathbf{v}_j' = \mathbf{R}_k \delta_{kj}. \tag{2.4}$$

The case where the dynamic and/or measurement equations are nonlinear is discussed in Section 4.

It is assumed that a rule of validation of the 'candidate measurements' is available such that it guarantees that the correct return will be retained with a given probability. This will be discussed in more detail later.

The model for the incorrect returns is as follows: Given that a certain number of these returns has been observed, their spatial distribution is assumed uniform and independent. The probability distribution of the number of incorrect returns is assumed to be a 'diffuse' uniform one, i.e. there is only a vague knowledge about it. This is discussed in more detail in the next section and in Appendix C.

Denote the set of validated measurements at time k as

$$Z_k = \{\mathbf{z}_{k,i}\}_{i=1}^{m_k} \tag{2.5}$$

and

$$Z^k \triangleq \{Z_j\}_{j=1}^{k}. \tag{2.6}$$

The minimum variance estimate, conditional mean, is, therefore

$$\hat{\mathbf{x}}_{k|k} = \int \mathbf{x}_k p(\mathbf{x}_k | Z^k) \, d\mathbf{x}_k. \tag{2.7}$$

The above equation constitutes the mathematical basis for the new approach: the best estimate, in the m.m.s.e. sense, is to be computed by conditioning upon all the observed random variables that, with some non-zero probability are dependent on $\mathbf{x}_k$. Therefore one has to use *all* the measurements that might have originated from the object in track. This conditioning on all the random variables that, with some nonzero probability, depend upon $\mathbf{x}_k$ is the main feature that distinguishes this approach from the previous ones.

3. THE PROBABILISTIC DATA ASSOCIATION (PDA) METHOD

Using the notations introduced in the previous section, define the following events

$$\chi_{k,i} = \{z_{k,i} \text{ is the correct return}\}, \quad i = 1, \ldots, m_k \quad (3.1)$$

and

$$\chi_{k,0} = \{\text{none of the validated returns is correct}\}. \quad (3.2)$$

Since only one return can be correct, the above events are mutually exclusive and exhaustive. With this, (2.7) can be written as follows:

$$\hat{\mathbf{x}}_{k|k} \triangleq E\{\mathbf{x}_k | Z^k\} = \sum_{i=0}^{m_k} E\{\mathbf{x}_k | \chi_{k,i}, Z^k\} P\{\chi_{k,i} | Z^k\}. \quad (3.3)^*$$

Therefore, what one needs is to find the expressions of

$$\beta_{k,i} \triangleq P\{\chi_{k,i} | Z^k\}, \quad i = 0, 1, \ldots, m_k, \quad (3.4)$$

i.e. the *a posteriori* probability of each return having originated from the object in track. This is the 'probabilistic data association'. As mentioned in the Introduction, it is assumed that no inference on the number of incorrect returns can be made from past data. It is also realistic to assume that the sensor 'looks' only within the validation region and, therefore, there is no information about the present 'clutter density'. Singer *et al.* [7] assume that this density is available and use a Poisson model with a fixed known parameter for the extraneous reports. While such a model fits the false alarms, it is questionable whether it is appropriate for clutter that might be, e.g. concentrated in certain regions of the space. Our assumption of ignorance, i.e. that no inference can be made on the number of incorrect returns, is shown in Appendix C to be equivalent to an improper or 'diffuse' uniform distribution on the number of incorrect returns.

The following assumptions are made in this PDA method

* The number of observed returns m_k is subsumed in the conditioning.

(1) The probability density of a measurement, conditioned upon past data and given that is correct,

$$p(z_{k,i} | \chi_{k,i}, Z^{k-1}) \triangleq f(z_{k,i} | Z^{k-1}) \quad (3.5)$$

it is assumed to be available. This density will be discussed more in the next section, where the modified filter is presented. In order to obtain a fixed-memory filter, this density will be approximated as normal.

(2) The density of a measurement given that it is incorrect is uniform in the validation region whose volume is denoted by V_k, i.e.

$$p(z_{k,i} | \chi_{k,j}, Z^{k-1}) = V_k^{-1}, \quad j \neq i. \quad (3.6)$$

(3) No inference can be made on the number of validated returns from past data.

(4) The probability of each return being correct, conditioned on the past data, is the same, i.e. no target signature information is used; the procedure can be modified to use such a type of information.

The derivation of the probabilities $\beta_{k,i}$ defined in (3.4) using Bayes' rule is carried out in detail in Appendix C. The *a posteriori* probability that the ith return is correct is obtained as

$$\beta_{k,i} = \left\{ f(z_{k,i} | Z^{k-1}) \middle/ \left[b_k + \sum_{i=1}^{m_k} f(z_{k,i} | Z^{k-1}) \right] \right\}, \quad i = 1, \ldots, m_k, \quad (3.7)$$

where

$$b_k \triangleq m_k V_k^{-1} \frac{\alpha_1 + \alpha_2 - \alpha_1 \alpha_2}{(1-\alpha_1)(1-\alpha_2)}. \quad (3.8)$$

α_1 is the probability that the correct return will not lie in the validation region and α_2 is the probability that the correct return will not be detected. The *a posteriori* probability that none of the returns is correct is

$$\beta_{k,0} = b_k \middle/ \left[b_k + \sum_{i=1}^{m_k} f(z_{k,i} | Z^{k-1}) \right]. \quad (3.9)$$

Equations (3.7) and (3.9) form the probabilistic data association method.

4. THE PROBABILISTIC-DATA-ASSOCIATION FILTER

As pointed out in [5], the computation of the exact probability density of the state conditioned upon all the validated measurements, which is a sum of Gaussian densities, is equivalent to track splitting and, consequently, too costly to be feasible. Therefore, an approximation has to be made in order to have an algorithm that can be implemented in real time.

To obtain such a filter, we shall approximate the density of the state conditioned upon the past observations as being normal with mean $\hat{\mathbf{x}}_{k|k-1}$ and covariance $\mathbf{P}_{k|k-1}$, i.e.

$$p(\mathbf{x}_k | Z^{k-1}) = \mathcal{N}(\hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}). \quad (4.1)$$

Similar assumptions were made in other related problems, e.g. in [8], in order to get an implementable solution. In [3] this assumption was also used, even though it was not explicitly stated. The practicality of assumption (4.1) is believed to be justified by the simplicity of the resulting algorithm and its performance in a practical problem as illustrated in the next section.

As noted earlier, because of computational considerations only those measurement returns which pass a certain validation test are considered for updating a particular track. Following (4.1) the residual, or innovation [9], corresponding to the correct return, denoted here by $\mathbf{z}_{k,j}$, is

$$\mathbf{v}_{k,j} \triangleq \mathbf{z}_{k,j} - \hat{\mathbf{z}}_{k|k-1}, \quad (4.2)$$

where $\hat{\mathbf{z}}_{k|k-1}$ is the conditional mean of the observation based on approximation (4.1), will also be assumed to be normally distributed with mean zero and covariance $\mathbf{S}_k$, given by

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k' + \mathbf{R}_k. \quad (4.3)$$

In view of this, one can write the validation test as follows: accept the measurements at time k that satisfy

$$\rho_{k,i} \triangleq \rho_k(\mathbf{v}_{k,i}) \leq \gamma, \quad (4.4)$$

where γ is a threshold,

$$\rho_k(\mathbf{v}) \triangleq \mathbf{v}' \mathbf{S}_k^{-1} \mathbf{v} \quad (4.5)$$

and $\mathbf{v}_{k,i}$ denotes the innovation corresponding to the measurement $\mathbf{z}_{k,i}$. The value $\rho_k(\mathbf{v})$ defined in (4.5) shall be referred to from now on as distance or norm squared.

The test (4.4) actually represents the requirement that a valid measurement be in the ellipsoid of a given probability concentration, a 'confidence ellipsoid' [10]. Namely, the constant γ is obtained by noticing that for the correct return, the distance ρ has an r degrees of freedom χ^2 distribution. This is done using χ^2 tables for a given probability α_1 of rejecting the correct return.

The approximate sufficient statistic of the past measurements according to (4.1) is denoted as

$$Y_{k|k-1} = \{\hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}\}. \quad (4.6)$$

Following (4.1)–(4.5) the probability density (3.5) of a measurement given that it originated from the object in track and it has been validated, conditioned upon Z^{k-1} is a truncated normal density, i.e.

$$f(\mathbf{z}_{k,i} | Y_{k|k-1}) = (1-\alpha_1)^{-1} \mathcal{N}(H_k \hat{\mathbf{x}}_{k|k-1}, \mathbf{S}_k) \quad (4.7)$$

inside the validation region and zero outside it.

With assumption (4.1) and the PDA method described in the previous section, the probabilistic data association filter (PDAF) is derived in Appendix A. The approximate conditional mean of the state is obtained as

$$E\{\mathbf{x}_k | Z_k, Y_{k|k-1}\} \triangleq \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{W}_k \mathbf{v}_k, \quad (4.8)$$

where

$$\mathbf{v}_k \triangleq \sum_{i=1}^{m_k} \beta_{k,i} \mathbf{v}_{k,i} \quad (4.9)$$

is the weighted innovation which uses all the validated measurements. The estimate $\hat{\mathbf{x}}_{k|k}$ is a nonlinear function of the observations via the coefficients $\beta_{k,i}$. The coefficient $\beta_{k,i}$ is the *a posteriori* probability that $\mathbf{z}_{k,i}$ originated from the object in track and $\mathbf{W}_k$ is the weighting matrix given by

$$\mathbf{W}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k' \mathbf{S}_k^{-1}. \quad (4.10)$$

The covariance associated with the estimate (4.9) is

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k}^0 + \mathbf{P}_k, \quad (4.11)$$

where $\mathbf{P}_{k|k}^0$ is the covariance of the update if we have only one return and

$$\mathbf{P}_k \triangleq \mathbf{W}_k \left[\sum_{i=1}^{m_k} \beta_{k,i} \mathbf{v}_{k,i} \mathbf{v}_{k,i}' - \mathbf{v}_k \mathbf{v}_k' \right] \mathbf{W}_k' \quad (4.12)$$

is a positive semidefinite matrix which shows the effect of the incorrect measurements by increasing the covariance of the update $\mathbf{P}_{k|k}$ as described in Appendix A. As can be seen from (4.12) the confidence on the estimate is a function of the actual number of validated returns and their location. This is an important feature of the new filter.

Another important feature of this algorithm is that its computational requirements are the same as the standard filter's when only one return falls in the window and increase only when the need of processing multiple returns arises. Based upon a computer instruction count, Table 1 shows the approximate increase of the computational requirements of the new filter compared to the standard filter* as a function of the number of returns in the validation region.

TABLE 1. COMPUTATIONAL REQUIREMENTS OF THE PDAF VS. THE STANDARD FILTER

Number of validated returns	1	2	3	4
Factor of increase	1	1·4	1·6	1·8

In Fig. 1 the expected computational requirements, per sampling time, relative to the standard filter, of the PDAF are shown vs those of the track-split filter (TSF). As can be seen, the requirements of the TSF grow exponentially with time when $\bar{r}$, the expected number of incorrect returns per

* The standard filter utilizes only one return, the nearest one, as if it were the correct one, i.e. the covariance and gain are not modified.

window is greater than 1; Appendix B discusses the pertinent equations. While these numbers are an upper bound because some of the tracks might be dropped, the computer might still saturate rapidly. As will be illustrated in the next section, the PDAF can be used successfully in a certain range of $\bar{r}$ where the TSF is not feasible and its performance is superior to the standard filter's performance.

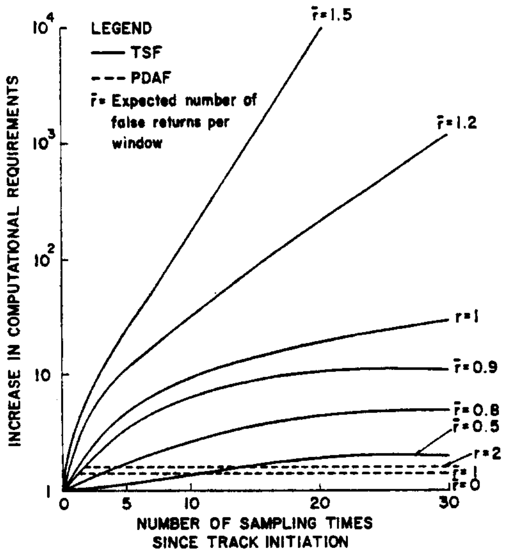

FIG. 1. Comparison of computational requirement of TSF and PDAF. (Normalized w.r.t. the standard filter.)

In problems where the dynamics and/or the measurements are nonlinear, the PDA method using the assumptions of Section 3 can be used with a filter as described by equations (4.8)–(4.12) based on the linearized system about the latest estimate. The additional errors introduced by the linearization can be handled in a manner similar to the one used for extended Kalman filters, e.g. by increasing the process noise covariance [11]. Simulations results using such an 'extended' PDAF for a practical problem are presented in the next section.

5. SIMULATION RESULTS

The problem on which this new algorithm was simulated is the tracking of a target with a seven-dimensional state with two angle-only sensors. A relatively complex and nonlinear example was chosen to assess the performance of the PDAF. Since such a simulation is more realistic, the conclusions based on it are of more practical significance than those that could be drawn from a simple example. The additional difficulty arising from the nonlinearities was solved by using as a standard filter the extended Kalman filter with the linearization performed about the latest estimate. This filter was using the 'nearest neighbor' as if it were the correct return, as it is done in most applications. The PDAF consisted of equations (4.8)–(4.12), where the gain and covariance were evaluated using linearization about the latest estimate, combined with the PDA method as described in Section 3.

The dynamic equations of the target were, in inertial coordinates,

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t)] + \mathbf{w}(t), \tag{5.1}$$

where w is the process noise,

$$\mathbf{x} = [x, y, z, \dot{x}, \dot{y}, \dot{z}, \alpha]', \tag{5.2}$$

$$f(\mathbf{x}) = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \alpha\dfrac{\dot{x}+\Omega y}{v_R} - \mu\dfrac{x}{(x^2+y^2+z^2)^{3/2}} \\ \alpha\dfrac{\dot{y}-\Omega x}{v_R} - \mu\dfrac{y}{(x^2+y^2+z^2)^{3/2}} \\ \alpha\dfrac{\dot{z}}{v_R} - \mu\dfrac{z}{(x^2+y^2+z^2)^{3/2}} \\ \dfrac{\alpha^2}{c} \end{bmatrix}, \tag{5.3}$$

x, y, z are the Cartesian position coordinates, α is the net difference between thrust and drag, both axial, per unit mass

$$\alpha = (T-D)/m \tag{5.4}$$

which satisfies the differential equation

$$\dot{\alpha} = \alpha^2/c, \tag{5.5}$$

where $c = 5$ km/sec and $\alpha(0) = 20$ m/sec^2. Ω denotes the spin rate of the earth,

$$v_R{}^2 = (\dot{x}+\Omega y)^2 + (\dot{y}-\Omega x)^2 + (\dot{z})^2 \tag{5.6}$$

and μ is the gravitation constant. The target trajectory was generated without noise; however, the filter, in order to account for the nonlinearities of the system, assumed a process noise as in (5.1) with a 'tuned' covariance equal to a fraction (10^{-2}) of the covariance of the updated state.

Each sensor measures the azimuth and elevation angles of the line of sight to the target with a variance of 10^{-8} rad^2. These sensors are located at different points and this guarantees the complete observability of the target's trajectory. In a 'clean' environment the extended Kalman filter reached steady state after about ten observations and the resulting 99 per cent validation region, the 'standard window', was used as the basis to define the clutter density. In the simulation the incorrect returns were generated using a random number generator as described in Appendix D. The parameter of the study was $\bar{r}$, the expected number of undesired returns in the standard window.

As pointed out earlier, the PDAF does not use any environmental parameter and, therefore, can be used with no modifications in any environment, even if $\bar{r}$ is time-varying.

The purpose of the simulation was to investigate the tracking capability of the PDAF vs the standard filters. At this stage, no theoretical model of the probability of losing track with the new filter is available. This is due to the fact that the covariances are computable in real time since they depend upon the number of validated measurements. The tracking capability of the PDAF has been, therefore, studied via simulations.

The length of each run was 100 samples, for each of the two sensors. Figure 2 presents the

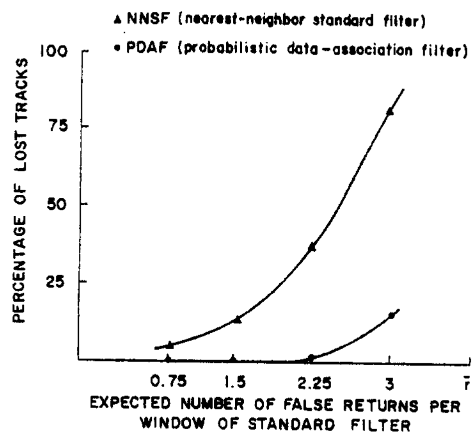

FIG. 2. Comparison of tracking capability of NNSF vs PDAF.

percentage of lost tracks for the NNSF compared to the PDAF for various clutter densities. These numbers were obtained from 50 Monte Carlo runs; a track was considered lost when the correct measurement was not in the validation region of at least one of the sensors for at least the last 20 sampling times. This is a reasonable definition of a lost track because in this case the final errors were very large. As can be seen, the PDAF is able to reliably track even for $\bar{r} = 2$ while for $\bar{r} = 0{\cdot}75$ the NNSF already has approximately a 4% probability of losing track.* Also note that for $\bar{r} > 1$ the track splitting filter (TSF) computational requirements increase exponentially with time (see Fig. 1). Therefore, the PDAF appears to be useful in the region where the NNSF has already a high probability of losing track and the TSF is unfeasible.

Figures 3 and 4 show the average position and velocity estimation errors respectively, from 50 runs, of the PDAF for two values of $\bar{r}$, compared to the average error in clean environment. For $\bar{r} = 0{\cdot}75$ the performance of the PDAF is the same as in a clean environment and is not shown. As can be seen, for $\bar{r} = 1{\cdot}5$ there is only a slight degradation in the performance and even for $\bar{r} = 2{\cdot}25$ it is still not very large. On the other hand, the NNSF, at $\bar{r} = 2{\cdot}25$ has a probability of losing track of approximately 40 per cent as indicated in Fig. 2, and compared to this the PDAF's performance is very good.

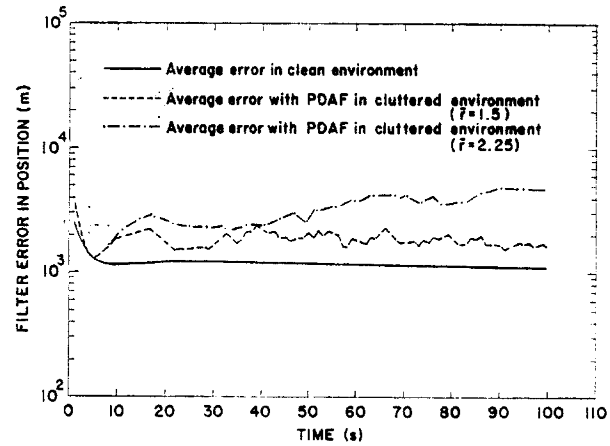

FIG. 3. Position estimation errors in clean and cluttered environments.

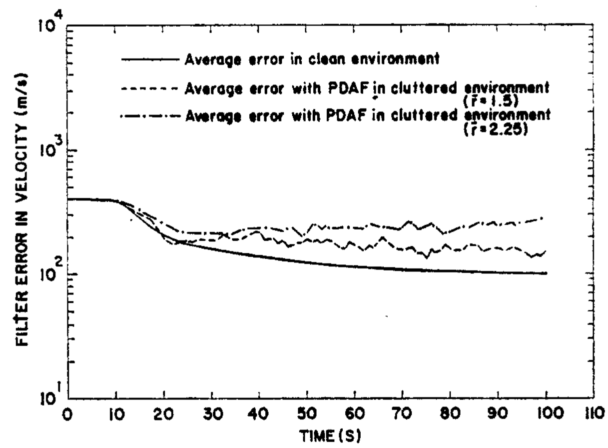

FIG. 4. Velocity estimation errors in clean and cluttered environments.

Therefore, as illustrated by this simulation, the PDAF can extend the tracking capability into the region of high clutter density where the NNSF becomes unreliable because of its high probability of losing track.

6. CONCLUSION

A significant improvement can be obtained in the performance of a filter operating in a cluttered environment when it utilizes all the measurements that pass a certain validation test. The procedure to utilize these measurements is to compute the *a posteriori* probability of each measurement having

---

* Note that the expected number of clutter returns in the window of the PDAF is greater than $\bar{r}$ because the window of the PDAF is increased according to (4.12).

originated from the object in track 'probabilistic data association') and then incorporating them into an appropriately modified filter, the PDAF. In order to obtain an algorithm that has computational requirements suitable for real-time application, a number of approximations were made. The PDAF, as developed here, is suitable only for nonmaneuvering targets. Its performance was illustrated by a simulation in which the PDAF was compared with the standard filter currently in use for this type of problem.

Acknowledgements—The authors wish to thank Dr. R. E. Larson for encouragement during this work, Professor M. Athans and Dr A. G. Jaffer for stimulating discussions on this subject and Ms. N. Rosenfield for the excellent programming support. Special thanks are due to the Editor and reviewers for their constructive criticism.

REFERENCES

[1] R. K. MEHRA: A comparison of several non-linear filters for reentry vehicle tracking. *IEEE Trans. Aut. Control* **AC-16**, 307–319 (1971).

[2] R. G. SEA: An efficient suboptimal decision procedure for associating sensor data with stored tracks in real-time surveillance systems. *Proc. of the 1971 IEEE Conference on Decision and Control*, pp. 33–37, Miami Beach, Fla., December (1971).

[3] R. A. SINGER and J. J. STEIN: An optimal tracking filter for processing sensor data of imprecisely determined origin in surveillance systems. *Proc. of the 1971 IEEE Conference on Decision and Control*, pp. 171–175, Miami Beach, Fla., December (1971).

[4] R. A. SINGER and R. SEA: New results in optimizing surveillance system tracking and data correlation performance in dense multitarget environments. *IEEE Trans. Aut. Control* **AC-18**, 571–582 (1973).

[5] Y. BAR-SHALOM and A. JAFFER: Adaptive nonlinear filtering for tracking with measurements of uncertain origin. *Proc. of the 1972 IEEE Conference on Decision and Control*, pp. 243–247, New Orleans, La., December (1972).

[6] A. G. JAFFER and Y. BAR-SHALOM: On optimal tracking in multiple-target environments. *Proceedings of the Third Symposium on Nonlinear Estimation Theory and its Applications*, pp. 112–117, San Diego, California, September (1972).

[7] R. SINGER, R. SEA and K. HOUSEWRIGHT: Derivation and evaluation of improved tracking filters for use in dense multitarget environments. *1973 Symp. on Inf. Theory*.

[8] G. A. ACKERSON and K. S. FU: On state estimation in switching environments. *IEEE Trans. Aut. Control* **AC-15**, 10–17 (1970).

[9] T. KAILATH: An innovations approach to least-squares estimation. Part I: Linear filtering in additive white noise. *IEEE Trans. Aut. Control* **AC-13**, 646–654 (1968).

[10] T. S. FERGUSON: *Mathematical Statistics*. Academic Press, New York (1967).

[11] A. JAZWINSKI: *Stochastic Processes and Filtering Theory*. Academic Press, New York (1970).

[12] M. DEGROOT: *Optimal Statistical Decisions*. McGraw-Hill (1970).

[13] Y. BAR-SHALOM and E. TSE: Tracking in a cluttered environment with probabilistic data association. *Proc. of the Fourth Symp. on Nonlinear Estimation, Theory and its Applications*, pp. 13–22, San Diego, California, September (1973).

[14] A. PAPOULIS: *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York (1965).

APPENDIX A

DERIVATION OF THE PDA FILTER

Noting that the events $\chi_{k,i}$, $i = 0, \ldots, m_k$, defined in (3.1) and (3.2) are mutually exclusive and exhaustive, one has

$$p(\mathbf{x}_k | Z_k, Y_{k|k-1}) = \sum_{i=1}^{m_k} p(\mathbf{x}_k | \chi_{k,i}, Z_k, Y_{k|k-1}) \beta_{k,i}, \quad \text{(A.1)}$$

where

$$\beta_{k,i} \triangleq P\{\chi_{k,i} | Z_k, Y_{k|k-1}\}. \quad \text{(A.2)}$$

From the definition of $\chi_{k,i}$ and (4.1) it is easy to see that, for $i = 1, \ldots, m_k$

$$p(\mathbf{x}_k | \chi_{k,i}, Z_k, Y_{k|k-1}) = p(\mathbf{x}_k | \chi_{k\,i}, \mathbf{z}_{k,i}, Y_{k|k-1}) = \mathcal{N}(\hat{\mathbf{x}}_{k|k,i}, \mathbf{P}_{k|k,i}), \quad \text{(A.3)}$$

where

$$\hat{\mathbf{x}}_{k|k,i} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{W}_{k,i} \boldsymbol{\nu}_{k,i}, \quad \text{(A.4)}$$

$$\mathbf{W}_{k,i} = \mathbf{P}_{k|k-1} \mathbf{H}_k' [\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k' + \mathbf{R}_k]^{-1} \triangleq \mathbf{W}_k, \quad \text{(A.5)}$$

$$\mathbf{P}_{k|k,i} = (I - \mathbf{W}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} \triangleq \mathbf{P}_{k|k}^*. \quad \text{(A.6)}$$

Note that the weighting matrix $\mathbf{W}_{k,i}$ is independent of i. However, $\mathbf{P}_{k|k,i} = \mathbf{P}_{k|k}^*$ in (A.6) is the covariance of the estimate (A.4) conditioned upon $\chi_{k,i}$ and even though it is independent of the particular value of i, it is not the covariance of the final estimate as will be seen later.

If the event $\chi_{k,0}$ occurs, i.e. none of the validated returns is correct, then (A.3) and (A.6) hold and

$$\mathbf{W}_{k,0} = \mathbf{0}, \quad \text{(A.5')}$$

$$\mathbf{P}_{k|k,0} = \mathbf{P}_{k|k-1}, \quad \text{(A.6')}$$

Therefore one now has the densities that enter into the summation in (A.1). Note that they are all normal.

The conditional mean of the state is readily obtained from (A.1) and can be written as

$$E\{\mathbf{x}_k | Z_k, Y_{k|k-1}\} \triangleq \hat{\mathbf{x}}_{k|k} = \sum_{i=0}^{m_k} \hat{\mathbf{x}}_{k|k\,i} \beta_{k,i} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{W}_k \boldsymbol{\nu}_k, \quad \text{(A.7)}$$

where

$$\boldsymbol{\nu}_k \triangleq \sum_{i=1}^{m_k} \beta_{k,i} \boldsymbol{\nu}_{k,i} \quad \text{(A.8)}$$

Notice that the estimate $\hat{\mathbf{x}}_{k|k}$ is a nonlinear function of the observations via the coefficients $\beta_{k,i}$.

The variance associated with the above estimate is obtained as follows

$$\mathbf{P}_{k|k} = \int [\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}] [\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}]' p(\mathbf{x}_k | Z_k, Y_{k|k-1}) \, \mathrm{d}\mathbf{x}_k. \quad \text{(A.9)}$$

Using the total probability theorem [14] and notation (A.2), the above can be written as

$$\mathbf{P}_{k|k} = \sum_{i=0}^{m_k} \beta_{k,i} \int [\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}][\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}]' p(\mathbf{x}_k \mid \chi_{k,i}, Z_k, Y_{k|k-1})\, \mathrm{d}\mathbf{x}_k. \tag{A.10}$$

After some algebra one obtains

$$P_{k|k} = \mathbf{P}_{k|k}{}^0 + \sum_{i=0}^{m_k} \beta_{k,i} \hat{\mathbf{x}}_{k|k,i} \hat{\mathbf{x}}_{k|k,i}{}' - \hat{\mathbf{x}}_{k|k} \hat{\mathbf{x}}_{k|k}{}'$$

$$= \mathbf{P}_{k|k}{}^0 + \mathbf{W}_k \left[\sum_{i=1}^{m_k} \beta_{k,i} \mathbf{v}_{k,i} \mathbf{v}_{k,i}{}' - \mathbf{v}_k \mathbf{v}_k{}' \right] \mathbf{W}_k{}'$$

$$\triangleq \mathbf{P}_{k|k}{}^0 + \dot{\mathbf{P}}_k. \tag{A.11}$$

The adaptivity of this filter can be readily seen from (A.14): the confidence on the estimate is a function of the actual number of validated returns. The matrix $\mathbf{P}_k$ shows the effect of the measurements that did not originate from the object in track by increasing the covariance of the estimate. This follows from the fact that $\mathbf{P}_k$ is positive semi definite as shown in [13].

Note that the computation of the covariance requires real-time data. This is a general characteristic of nonlinear filters.

APPENDIX B
COMPUTATIONAL REQUIREMENTS FOR THE TRACK-SPLITTING FILTER (TSF)

Let the number of incorrect returns detected in the ith window at time k be $r_i(k)$, a random variable. Starting from time $k = 0$, the number of windows that are set up at a particular time k, if the track is split every time more than one return is detected, is obtained as follows.

At $k = 0$ there is one window in which

$$1\ (\text{correct}) + R(0)\ (\text{incorrect})$$

returns are detected. Corresponding to each of these returns, a new window is set up at $k = 1$. In one of these windows there are $1 + r_0(1)$ returns, including the correct one, while in the other there are $r_i(1)$, $i = 1, \ldots, R(0)$ returns. The subscript zero denotes the window in which the correct return lies. The total number of returns, assuming the windows do not overlap, is now $1 + R(1)$ where

$$R(1) = \sum_{i=0}^{R(0)} r_i(1). \tag{B.1}$$

Similarly, at time $k = 2$ there are $1 + R(2)$ returns where

$$R(2) = \sum_{i=0}^{R(1)} r_i(2). \tag{B.2}$$

In general, the number of returns at time k is $1 + R(k)$, where

$$R(k) = \sum_{i=0}^{R(k-1)} r_i(k). \tag{B.3}$$

Assume that the expected number of incorrect returns in a gate is

$$Er_i(k) = \bar{r}(k) \tag{B.4}$$

and that $r_i(k)$ is independent of $r_j(m)$ for $k \neq m$ and $\forall\, i,j$. Then, the expected number of tracks at time is

$$\bar{N}(k) = 1 + ER(k) \tag{B.5}$$

and

$$ER(k) = \bar{r}(k)[1 + ER(k-1)]$$

$$= \bar{r}(k)\{1 + \bar{r}(k-1)[1 + ER(k-2)]\}$$

$$= \sum_{j=1}^{k} \prod_{l=j}^{k} \bar{r}(l). \tag{B.6}$$

Let α be the amount of computations per track. Then the expected computational requirements at time k are, assuming for simplicity $\bar{r}(k) = \bar{r}$

$$\bar{C}(k) = \alpha \bar{N}(k) = \alpha \left[1 + \sum_{j=1}^{k} \bar{r}^j \right]$$

$$= \begin{cases} \alpha \left[1 + \dfrac{\bar{r} - \bar{r}^{k+1}}{1 - \bar{r}} \right], & \bar{r} \neq 1 \\ \alpha(1 + k), & \bar{r} = 1. \end{cases} \tag{B.7}$$

If $\bar{r}(k)$ is time varying then the closed form expression (B.7) does not hold and (B.6) has to be evaluated directly. Figure 4 shows the increase with time of the expected computational requirements for various values of $\bar{r}$. It can be seen that if $\bar{r} < 1$, i.e. expected number of incorrect returns per window is less than unity, $\bar{C}$ tends to a constant, namely,

$$\lim_{k \to \infty} \bar{C}(k) = \alpha \frac{1}{1 - \bar{r}}. \tag{B.8}$$

APPENDIX C
DERIVATION OF THE PDA METHOD

Using Bayes' rule, the probabilities from (3.4) can be written as [P1]

$$P\{\chi_{k,i} \mid Z^k, m_k\} = P\{\chi_{k,i} \mid Z_k, Z^{k-1}, m_k\}$$

$$= c_k{}^{-1} p(Z_k \mid \chi_{k,i}, Z^{k-1}, m_k) \times P\{m_k \mid \chi_{k,i}, Z^{k-1}\} P\{\chi_{k,i} \mid Z^{k-1}\}, \tag{C.1}$$

where the conditioning on the total number of observed returns at time k being equal to m_k has now been written out explicitly, and

$$c_k = \sum_{i=0}^{m_k} p(Z_k \mid \chi_{k,i}, Z^{k-1}, m_k) P\{m_k \mid \chi_{k,i}, Z^{k-1}\} \times P\{\chi_{k,i} \mid Z^{k-1}\} \tag{C.2}$$

is a normalization constant.

Note that the probability of observing a *total* number of returns equal to m_k conditioned upon $\chi_{k,0}$ ('none of the returns is correct') and Z^{k-1} is

$$P\{m_k | \chi_{k,0}, Z^{k-1}\} = P\{I_k = m_k | Z^{k-1}\}, \quad \text{(C.3)}$$

where I_k is the number of *incorrect* returns that are observed at time k. The probability of observing a total number of returns equal to m_k, conditioned upon $\chi_{k,i}$, $i \neq 0$, and Z^{k-1} is

$$P\{m_k | \chi_{k,i}, Z^{k-1}\} = P\{I_k = m_k - 1 | Z^{k-1}\}, \quad i = 1, \ldots, m_k \quad \text{(C.4)}$$

because in this case one of the returns is correct.

The assumption that no inference can be made on the number of incorrect returns from past data is translated into treating the events $I_k = m_k$ and $I_k = m_k - 1$ as equally likely, i.e.

$$P\{I_k = m_k | Z^{k-1}\} = P\{I_k = m_k - 1 | Z^{k-1}\} = \Pi_0. \quad \text{(C.5)}$$

Note that the value of Π_0 is irrelevant because it cancels out in (C.1). Such a distribution, which is improper because its integral is infinite, is often used to describe vague prior knowledge; however, due to the above-mentioned cancellation the resulting posterior *is proper*. A detailed discussion of improper prior distributions can be found in De Groot [12].

For $i = 0$, i.e. when all the validated returns are incorrect ones, their joint density is, in view of the discussion of Section 2,

$$p(Z_k | \chi_{k,0}, Z^{k-1}, m_k) = \prod_{i=1}^{m_k} p(z_{k,i} | \chi_{k,0}, Z^{k-1}, m_k) = V_k^{-m_k} \quad \text{(C.6)}$$

since, denoting by V_k the volume of the validation region,

$$p(z_{k,i} | \chi_{k,j}, Z^{k-1}, m_k) = V_k^{-1}, \quad j \neq i \quad \text{(C.7)}$$

because of our assumption on the incorrect measurements being uniformly distributed.

The probability of $\chi_{k,0}$ based on past data is

$$P\{\chi_{k,0} | Z^{k-1}\} = \alpha_1 + (1 - \alpha_1)\,\alpha_2, \quad \text{(C.8)}$$

where α_1 is the probability that the correct return will not lie in the validation region and α_2 is the probability that the correct return will not be detected.

For $i = 1, \ldots, m_k$ one can write the density on the right-hand side of (C.1) as

$$\begin{aligned} p\{Z_k | \chi_{k,i}, Z^{k-1}, m_k\} &= p(z_{k,i} | \chi_{k,i}, Z^{k-1}) \prod_{\substack{j=1 \\ j \neq i}}^{m_k} p(z_{k,j} | \chi_{k,i}, Z^{k-1}, m_k) \\ &= [V_k]^{-(m_k - 1)} p(z_{k,i} | \chi_{k,i}, Z^{k-1}). \end{aligned} \quad \text{(C.9)}$$

The density of $\mathbf{z}_{k,i}$, given that it is correct, conditioned upon the past data is denoted as

$$p(\mathbf{z}_{k,i} | \chi_{k,i}, Z^{k-1}) \triangleq f(\mathbf{z}_{k,i} | Z^{k-1}) \quad \text{(C.10)}$$

assumed to be known.

The probability of $\chi_{k,i}$ conditioned upon past data is assumed the same for all i, unless target signature information can be used. If no such information is available, then

$$P\{\chi_{k,i} | Z^{k-1}\} = \frac{1 - P\{\chi_{k,0} | Z^{k-1}\}}{m_k} = \frac{(1 - \alpha_1)(1 - \alpha_2)}{m_k}, \quad i = 1, \ldots, m_k. \quad \text{(C.11)}$$

Inserting (C.6), (C.9) and (C.11) into (C.1) yields the equations of the PDA method, (3.7) and (3.9).

APPENDIX D
SIMULATION OF INCORRECT RETURNS

The incorrect returns were generated, for the purpose of this simulation, as follows. Validation regions (99 per cent confidence ellipses) were set up at every time for each sensor. The area of the validation region, a window, of the standard filter in steady state is denoted by A_v^0. The parameter of our study is

$$\bar{r} = A_v^0 \bar{\rho}, \quad \text{(D.1)}$$

the expected number of incorrect returns in the standard window and $\bar{\rho}$ is their expected density. Then, with A_v being the actual area of the validation region, and time-varying for the PDAF as indicated by (4.12)

$$n = [10 A_v \bar{\rho} + 1] \quad \text{(D.2)*}$$

measurements were generated independently and

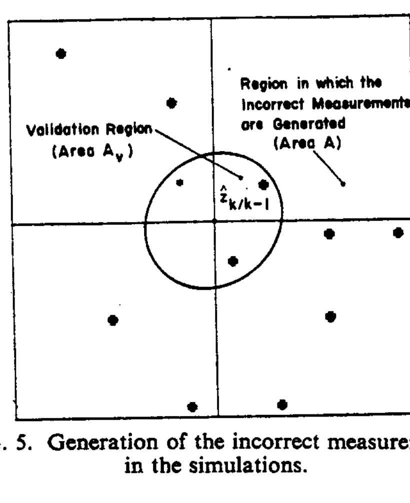

FIG. 5. Generation of the incorrect measurements in the simulations.

* $[x]$ is the largest integer smaller or equal to x.

uniformly distributed in a square, centered at the predicted location of the correct measurement, with an area

$$A = n/\bar{\rho} \tag{D.3}$$

This is illustrated in Fig. 5 where the correct measurement is denoted by '*' and the incorrect ones by '+'. In this way, since $A \gg A_v$, the actual number of validated incorrect returns was indeed random with mean $A_v \rho$. This procedure to generate the incorrect returns is believed to simulate quite accurately what might happen in reality when there is clutter or the false alarm rate is high. The parameter $\bar{r}$ was kept constant during the length of each run.